

A STATISTICAL IMAGE-BASED APPROACH FOR THE 3D RECONSTRUCTION OF THE SCOLIOTIC SPINE FROM BIPLANAR RADIOGRAPHS

S. Kadoury, F. Cheriet

Department of Biomedical Engineering
Ecole Polytechnique of Montréal, Canada

H. Labelle (MD)

CHU St-Justine Hospital Research Center
Montréal, Canada

ABSTRACT

In this paper, we propose a hybrid approach using a statistical 3D model of the spine generated from a database of 732 scoliotic patients with high-level anatomical primitives identified and matched on biplanar radiographic images for the three-dimensional reconstruction of the scoliotic spine. The 3D scoliotic curve reconstructed from a coronal and sagittal radiograph is used to generate an approximate statistical model based on a transformation algorithm which incorporates intuitive geometrical properties. An iterative optimization procedure integrating similarity measures such as deformable vertebral contours and epipolar constraints is then applied to globally refine the 3D anatomical landmarks on each vertebra level of the spine. A qualitative evaluation of the retro-projection of the vertebral contours obtained from the proposed method gave promising results while the quantitative comparison yield similar accuracy on the localization of low-level primitives such as the six landmarks identified by an expert on each vertebra.

Index Terms— 3D spine reconstruction, radiographs, statistical model, deformable contours, scoliosis.

1. INTRODUCTION

Several clinical studies in orthopedics have used 3D models of the spine for evaluating pathologies in spinal deformities such as idiopathic scoliosis. Biplanar radiographs are still the imaging modality which is most frequently used for the clinical assessment of spinal deformities since it allows the acquisition of patient's data in the natural standing posture. To generate a 3D model of the patient's spine, certain points (anatomical landmarks) on the vertebra within the image have to be located in order to obtain a three-dimensional model of the scoliotic spine using a triangulation algorithm [1]. Currently, this identification is performed manually by an expert operator and consists of locating six anatomical landmarks (2 endplate midpoints + 4 pedicle tips) on each vertebra from T1 (first thoracic vertebra) to L5 (last lumbar vertebra) on a coronal and sagittal radiograph. However, it is difficult to identify with precision low-level primitives

such as exact points and to match them accurately on a pair of views. Thus the repeatability of this procedure cannot be assured. Furthermore this task is a time-consuming, tedious and error-prone, and the quality of the 3D reconstruction is directly linked with the precision of 2D localization.

In order to reduce inaccuracy on the 2D localization of landmarks and to be a clinically useful procedure, studies were conducted to propose more automated methods. Statistical shape models, and more recently 2D-3D registration methods, have been the focus of a lot of attention for the 3D reconstruction of the human spine. While some have used preoperative computed tomography (CT) or magnetic resonance (MR) images to register with 2D radiographic images [2], others have used statistical *a priori* knowledge of the 3D geometric shapes in order to model the 2D vertebral shapes [3]. Currently most methods use statistical models capturing the geometrical knowledge of isolated scoliotic vertebrae [4] or ad hoc symbolic constraints on the whole spine shape [5]. A variability model (mean and dispersion) of the whole spine allowed increasing the accuracy of the 2D-3D registration algorithm [6]. However, none of these proposed methods have attempted to integrate a statistical model taking into account the set of admissible deformations for the whole scoliotic spine shape with high-level contour based information of the vertebrae extracted from the radiographic image. It is our belief that taking advantage of the full potential of the radiographic image content would increase the accuracy of the 3D reconstruction procedure.

The objective of this study is to propose a novel method for the 3D reconstruction of the spine by using the 3D spine centerline to predict an initial shape which reproduces the pathological deformations observed on a representative scoliotic spine database. Deformation of vertebral models, epipolar constraints and morphological/feature based information taken on the radiographs are then used to refine the 3D landmarks through a bundle adjustment approach.

2. MATERIAL AND METHODS

2.1. Reconstruction of a 3D spinal curve

To generate the patient-specific spinal curve in 3D, the spine centerline must first be extracted from the calibrated radio-

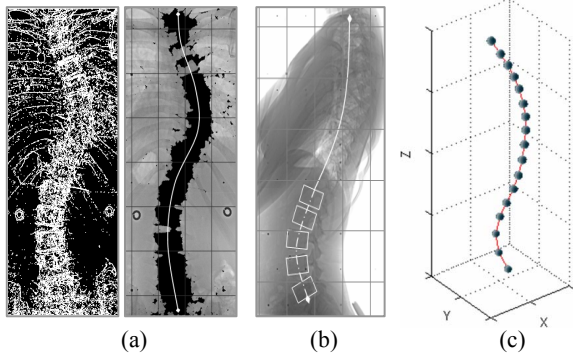


Fig. 1. (a) Coronal and (b) sagittal spine centerlines. (c) Corresponding 3D reconstruction of the spinal curve.

graphs to calculate the 3D coordinates of the curve using a triangulation algorithm. Both centerlines are obtained semi-automatically and modeled by 2D B-splines to ease the user's adjustments. A parametric 3D B-spline spinal curve $C_k(u) \in \mathbb{R}^3$ is then reconstructed as illustrated in Figure 1.

2.2. Approximate statistical 3D model of the spine

The 3D spinal curve $C_k(u)$ is used to predict an initial spine model from a 3D database containing 732 scoliotic spines demonstrating several types of deformities, by mapping the 3D curve to a low-dimensional subspace. We propose an algorithm derived from a locally linear embedding transformation [7], which is based on simple geometric intuitions to generate a personalized spine model. It computes a low-dimensional embedding of high-dimensionality data assumed to lie on a non-linear manifold, with the property that similar models of a spine in the high dimensional space remain nearby, and similarly remain co-located with respect to one another in the low dimensional space. Given N spine models expressed by the B-splines $C(u)_i$, $C(u)_i \in \mathbb{R}^D$, $i \in [1, N]$, each of dimensionality D , it provides N points Y_i , $Y_i \in \mathbb{R}^d$, $i \in [1, N]$ where $d \ll D$. The algorithm has four sequential steps:

Step 1. With an adequate number of data points available so that the underlying manifold can be considered to be “well-sampled” enough to represent the scoliotic population, each individual data point of the training set and its corresponding neighbors would be sufficiently close to lie within a locally linear patch on the manifold. The K closest neighbors are selected for each point using the Euclidean distance as a closeness measure.

Step 2. The second step involves solving for the manifold reconstruction *weights*. Clearly, the local geometry of the patches referred to in Step 1 can be described by linear coefficients that permit the reconstruction of every model point from knowledge of its neighbors. In order to determine the value of the weights, the reconstruction errors are measured by the cost function:

$$\varepsilon(W) = \sum_{i=1}^N \left\| C(u)_i - \sum_{j=1}^K W_{ij} C(u)_{ij} \right\|^2 \quad (1)$$

where $C(u)_i$ is a data vector and $\varepsilon(W)$ sums the squared distances between all data points and their corresponding reconstructed points. The weights W_{ij} represent the importance of the j^{th} data point to the reconstruction of the i^{th} element.

Step 3. The third step of the algorithm consists of mapping each high-dimensional $C(u)_i$ to a low-dimensional Y_i , representing the global internal coordinates using a cost function which minimizes the reconstruction error:

$$\Phi(Y) = \sum_{i=1}^N \left\| Y_i - \sum_{j=1}^K W_{ij} Y_{ij} \right\|^2 \quad (2)$$

The coordinates Y_i can be translated by a constant displacement without affecting the overall cost, $\Phi(Y)$. This degree of freedom is removed by requiring the coordinates to be centered at the origin, such that $\sum Y_i = 0$. The optimal embedding, up to a global rotation of the embedding space, is obtained from the bottom $d+1$ eigenvectors of the matrix M . The d eigenvectors form the d embedding coordinates.

Step 4. The final step applies an analytical method based on nonlinear regression to perform the inverse mapping from the d embedding. Given the original training data consisting of N (732) scoliotic spine models X_i , ($i=1, 2, \dots, N$), each of $D2$ dimension (output high-dim. space), and their respective projection Y_i (embedded data) obtained in step 3 of the algorithm for every $C(u)_i$ (computed from X_i), then each dimension of the $D2$ space can be regressed by:

$$X_{new} = F(Y_{new}) = [x_{new1}, \dots, x_{newD2}]^T = [f_1(Y_{new}), \dots, f_{D2}(Y_{new})]^T \quad (3)$$

where $x_i = f_i(Y) = \sum_j a_{ij} k(Y, Y_j) + b$ is a SVR regression model using a Radial Basis Function kernel, $X_{new} = (s_1, s_2, \dots, s_{17})$, where s_i is a vertebra model defined by $s_i = (p_1, p_2, \dots, p_6)$, and $p_i = (x_i, y_i, z_i)$ is a 3D vertebral landmark. Equation (3) provides a means of generating new spine models in $D2=102$ space ($D2=17$ vertebrae \times 6 landmarks) from a new embedded point and the training scoliotic data in the lower-dimensional d -space. This method not only allows restraining the search space for localizing the anatomical landmarks, but also avoids solving the point-matching problem between the biplane views. In the case that some landmarks are completely invisible, this method can offer an approximate position based on the statistical distribution of the pathological population.

2.3. Bundle adjustment of the 3D vertebral landmarks

The crude statistical 3D model of the personalized spine is subsequently refined by adjusting the 3D coordinates of the vertebrae. The set of 3D landmarks p_i for each vertebra s_i are globally adjusted based on the following measures.

Image gradient edge alignment: In order to integrate image-based information in the optimization process, we developed a similarity estimate based on the distance of the

projection of a 3D deformed model to the computed gradient of the radiographs. The approach would: 1) deform *prior* generic high resolution 3D vertebra model (17 in total) obtained from CT acquisitions, using a free form deformation technique with the set of landmarks p_i as the control points [1]; 2) project the triangulated mesh of the 3D model using the calibration matrices to create a silhouette onto the images; 3) compute a 2D distance map for these edges and; 4) sum over the distance map values at the locations indicated by the edges of the gradient image. Given the binary gradient radiographic image, the distance of an image point x to the projected edge structures $Y = \{y_i\}$ is $d(x) = \min_i |x - y_i|$. We can then express the proximity to edges by using a Gaussian expression controlled by σ^2 :

$$d(x) = \max_i \exp - \frac{(x - y_i)^2}{\sigma^2} \quad (4)$$

However due to the poor quality of the radiographic images, we do not have precise edge information and the gradient images may not correspond to the edge templates. We therefore define the proximity in (4) as:

$$d(x) = \sum_{j=1}^2 \sum_i p_{ji} \exp - \frac{(x - y_{ji})^2}{\sigma^2} \quad (5)$$

where p_{ji} is the probability for pixel y_i in image j of being an edge. To determine the values of p_{ji} , a two-dimensional proximity function $p(x)$ can be computed by convoluting the radiographic image with a large Gaussian kernel.

Epipolar geometry constraint: The calibration of the three-dimensional viewing geometry can also be used to constrain the landmark correspondence between the biplanar images. We therefore developed an iterative retro-projection method to help refine landmark position, by taking the current 3D landmark location, project it in 2D onto the coronal (PA)/ sagittal (SAG) views and measure the perpendicular distance of the projected coordinate on both views to its corresponding epipolar line. The distance error for the N landmark points is defined as:

$$\mathcal{E}(\xi) = \sum_{i=1}^N d(\hat{w}_i^{SAG}, \mathbf{F} \hat{w}_i^{PA})^2 + d(\hat{w}_i^{PA}, \mathbf{F}^T \hat{w}_i^{SAG})^2 \quad (6)$$

where $d(*)$ denotes the Euclidean distance of a point to a line, \hat{w}_i is the analytical projection of the 3D object point p_i obtained from standard perspective transformation formulae. $\mathbf{F} \hat{w}_i$ is the corresponding epipolar line on one image based on point p_i from the other image, and \mathbf{F} is the 4x4 fundamental matrix integrating the geometrical parameters ξ which describes the projective 3D structure of the scene.

Maximum Likelihood estimation: Finally, a maximum likelihood model estimation integrating 2D morphological and feature information was developed to measure the error given from the current data. This estimate expresses the measure of similarity between the current model points \hat{w}_i and an estimate $w_i(x)$ which encodes expert morphological

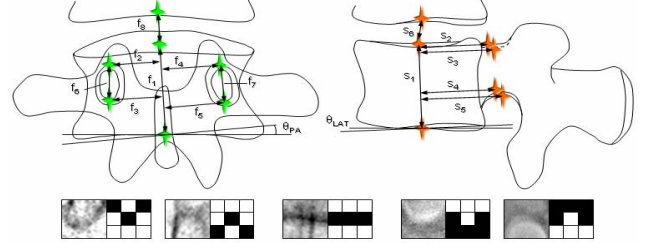


Fig. 2. Morphological estimation model (top) and landmark specific features (bottom).

knowledge of the relationships between the N ($N=6$) landmarks. Each landmark i is assigned to a specific function of $w_i(x)$ depending on the landmark type (i.e. pedicle tip), and is based on local vertebral height, width, orientation and relative distances between landmarks (Figure 2). The model also measures the similarity response of a rotation and scale invariant wavelet coefficient feature c_{msd} specific to the landmark type, at location \hat{w}_i on the image. The probability of this likelihood estimate is:

$$P \propto \prod_{i=1}^N \left\{ \exp \left[- \frac{1}{2} \left(\frac{\Psi_{PA}(\hat{w}_i^{PA}) + \Psi_{SAG}(\hat{w}_i^{SAG})}{2\sigma} \right)^2 \right] \Delta w \right\} \quad (7)$$

The similarity measures are defined as:

$$\Psi_{PA}(\hat{w}_i^{PA}) = (\hat{w}_i^{PA} - w_i^{PA}(x)) * c_{msd}(\hat{w}_i^{PA}) \quad (8)$$

$$\Psi_{SAG}(\hat{w}_i^{SAG}) = (\hat{w}_i^{SAG} - w_i^{SAG}(x)) * c_{msd}(\hat{w}_i^{SAG}) \quad (9)$$

where $w_i^{PA}(x) = \sum_j \delta_j f_j(x)$ and $w_i^{SAG}(x) = \sum_j \delta_j s_j(x)$ are the estimates on the coronal and sagittal plane respectively.

Cost function: For a bundle adjustment of the 3D landmark coordinates, a non-linear optimization method minimizes the cost function $E(s_l)$ with respect to all 6 anatomical landmarks in s_l at each vertebral level l (starting from L5 and progressing to T1), based on the measures taken on the biplane images. We used the Powell-Brent optimization method to minimize the following cost function combining equations (5), (6) and (7):

$$E(\bar{s}_l) = \omega_1 d(x) + \omega_2 \mathcal{E}(\xi) + \omega_3 / \log P \quad (10)$$

where $\bar{s}_l = R(\bar{s}_{l-1})[s_l] + T(\bar{s}_{l-1})$ takes into account the previous updated vertebra model, and (R, T) is the rigid displacement of landmarks p_i at the previous vertebra level s_{l-1} before/after optimization. Figure 3 summarizes the proposed scheme.

3. RESULTS AND DISCUSSION

To validate the proposed method, ten pairs of radiographic images from scoliotic patients treated at our clinical site were processed for personalized 3D reconstruction of their spine. The evolution of the cost function with respect to the estimate for any vertebral level. To assess the precision of the image similarity measure used for the optimization

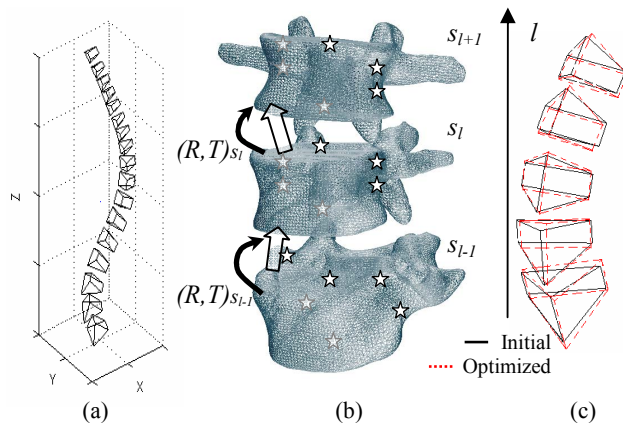


Fig. 3. (a) Approximate statistical 3D model. (b) Sequential refinement using deformable 3D prior models fitted on the radiographs (control points shown as stars). (c) Landmark adjustments brought by the optimization scheme.

procedure, Figure 4 shows promising results with the retro-projection of the deformed 3D vertebra contours (high-level primitive) fitting adequately to the bony edges of the corresponding vertebra in the coronal and sagittal radiographic image. Moreover, a quantitative evaluation showed the projected anatomical landmarks from the optimized 3D model yield similar accuracy to the gold standard 2D locations manually identified by a radiology expert on each vertebra. The overall 2D mean difference for the selected cases was of 2.3 ± 1.7 pixels, while the point to point mean difference between the 3D spine models issued from the proposed technique and from a manual identification yielded a 3D mean error of 1.8 ± 1.5 mm for lumbar vertebra and 2.2 ± 1.6 mm for thoracic vertebra.

One of the challenges in landmarking vertebrae on radiographs for generating 3D models is the poor visibility due to the superposition of several anatomical structures, specifically in the sagittal thoracic region of the spine. The human expert must therefore infer the landmark positions based only on his knowledge of the anatomical structure of the spine. The proposed method offers a more reliable approach to this problem by integrating statistical, image-based and morphological knowledge, and therefore becomes a suitable tool for clinical assessment of spinal deformities.

4. CONCLUSION

We presented a method to automatically compute a personalized 3D geometrical model of the spine based on the statistical distribution of a scoliotic population. Deformable 3D vertebral models projected on the image planes and epipolar/morphological constraints were used to refine the anatomical landmark coordinates of the model. We successfully applied our method to a group of scoliotic patients. Results presented in this paper suggest that an accurate geometry of the spine can be obtained by using a hybrid approach which captures a pathological population

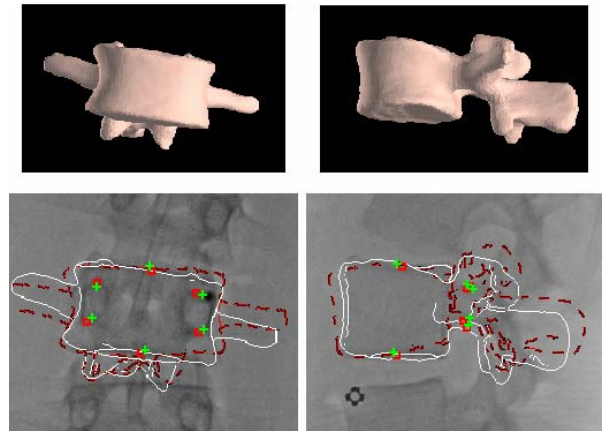


Fig. 4. Comparison of landmark and projected contours results from the manual (square/dashed line) and proposed method using 3D deformable models (cross/solid line).

exploits the image's content, offering an efficient method for spine 3D reconstruction in a routine clinical environment. We used the highest number of patients available ($N=732$) to create a predictive statistical model which would represent all types of deformity in the scoliotic population; however classifying the database into multiple pathological categories would reduce the size of N and offer a more representative sample of patients. Other future directions lies in increasing the accuracy of the method by modifying the objective function to integrate inter-vertebral variability models to obtain a better approximation of the spine and optimizing the deformable 3D models via a level-set approach. The proposed method can be extended to other medical reconstruction applications such as for the pelvis or femur, when a sufficient amount of prior data is available to adequately model various types pathologies.

5. REFERENCES

- [1] S. Delorme, Y. Petit, J. de Guise, *et al.* "Assessment of the 3-D reconstruction and high-resolution geometrical modeling of the human skeletal trunk from 2-D radiographic images," *IEEE Trans. Biomed. Engineering*, vol. 50, no. 8, 2003.
- [2] D. Tomazevic, B. Likar, T. Slivnik, and F. Pernus, "3D/2D Registration of CT and MR to X-Ray Images," *IEEE Trans. on Medical Imaging*, vol. 22, no. 11, 2003.
- [3] M. Fleute and S. Lavallée, "Nonrigid 3-D/2-D registration of images using a statistical model," in *MICCAI*, vol. 38, 1999.
- [4] S. Benamer, M. Mignotte, H. Labelle, and J. De Guise, "A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine," *IEEE Trans. Biomed. Engineering*, vol. 52, no. 12, 2005.
- [5] J. Novosad, F. Cheriet, *et al.*, "3D reconstruction of the spine from a single x-ray image and prior vertebrae models," *IEEE Trans. Biomed. Engineering*, vol. 51, no. 9, 2004.
- [6] J. Boisvert, X. Pennec, *et al.*, "3D anatomical variability assessment of the scoliotic spine using statistics on Lie groups," in *IEEE International Symposium on Biomedical Imaging*, 2006.
- [7] S. Roweis, and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, 2000.